Online learning First Sef
Supervised learning: Tooling to Usual Test
phose the phose the phose
On the learning: No such separation
Supervised/PAC (cornig: Distributional appropriate
Online learning: No such appropriate one made
Online learning: Thereaction
"Sequential", horizon: T" rounds
for
$$f=1, -- T$$

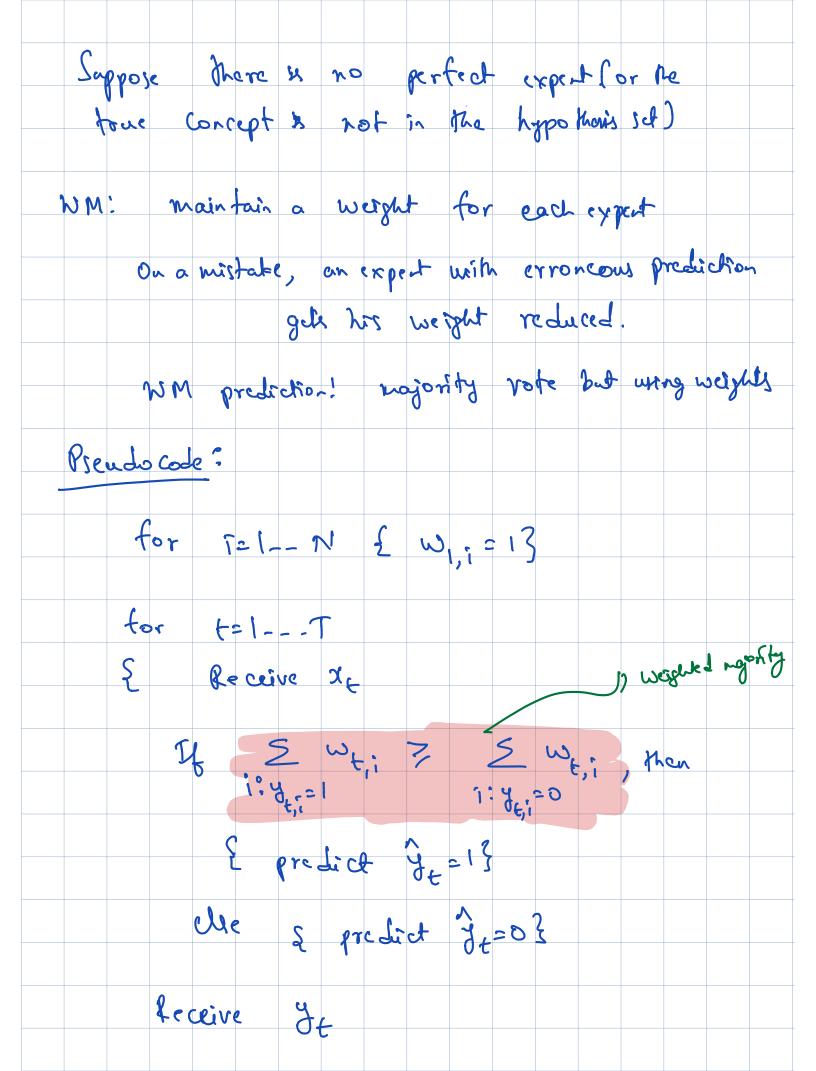
i Algorithmit receives $x_{t} \in X$
Supervise $Y_{t} \in Sincur loss L(S_{t}, S_{t})$
Supervise $Y_{t} \in Sincur loss L(S_{t}, S_{t})$
Supervise $Y_{t} \in Sincur loss L(S_{t}, S_{t})$

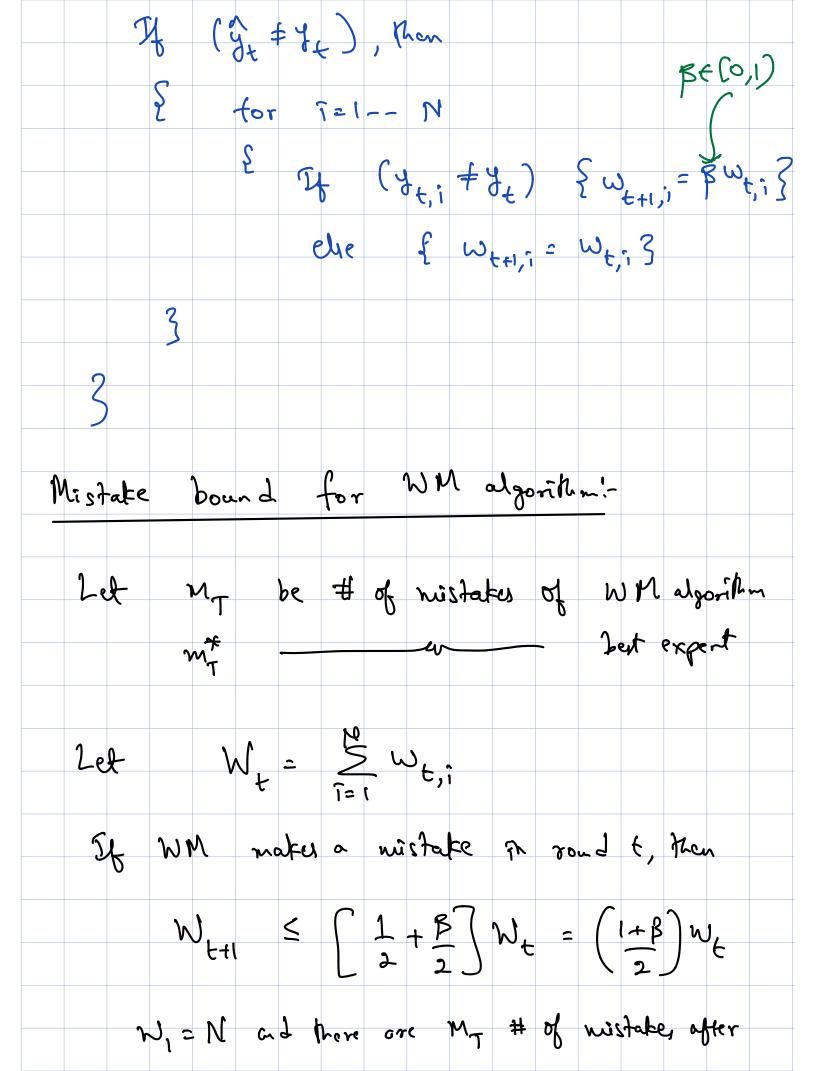
cum lative (oss $\sum L(\hat{y}_{t}, \hat{y}_{t})$ mininize セント Examples : 1) Manification problem. y={0,1} L(y, y') = [y - y'] $r L(y, y') = 1 \xi y \neq y' \xi$ 2 Regression $L(y, y') = (y - y')^2 \text{ with } Y \subseteq \mathbb{R}$ "Prediction with expert advice" In each round "t" algorithm A receives 2 E S & advice y E J, i=1,--N N = # of caperts alg & predicts & 4 observes L(Jr, Je) God : Mininize coulative loss_ Practical motivation !-

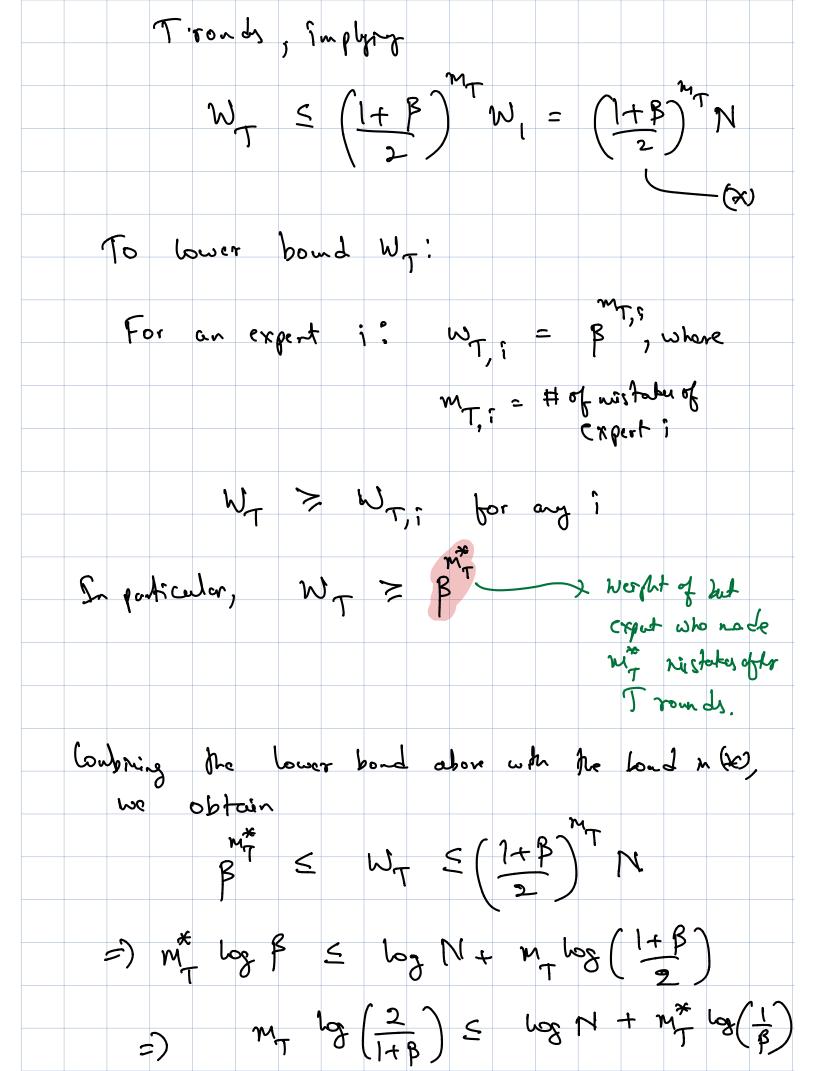
Movie recommendation Experti: IMDB, AT, your friend For a movie, each expert maked a prediction " Lod" to "Loop" Notion of regart : (Cunlative) Loss of expert i: $\sum_{t=1}^{T} L(y_{t,i}, y_t)$ u best expert : $\min_{f=1-N} \frac{t}{f_{2}} L(y_{t,1}, y_{t})$ $u = of alg f: \sum_{t=1}^{T} L(J_t, J_t)$ Regret $R_T = \sum_{t=1}^{T} L(\hat{y}_t, \hat{y}_t) - \min_{t=1}^{T} \sum_{t=1}^{T} L(\hat{y}_t, \hat{y}_t)$ Realizable core or the Core with the "perfect" expert: Loss: 0-1 Loss Binary clauification Want an algorithm with the Smallest ## of mytakes To judge an algorithm A, we we the max # of

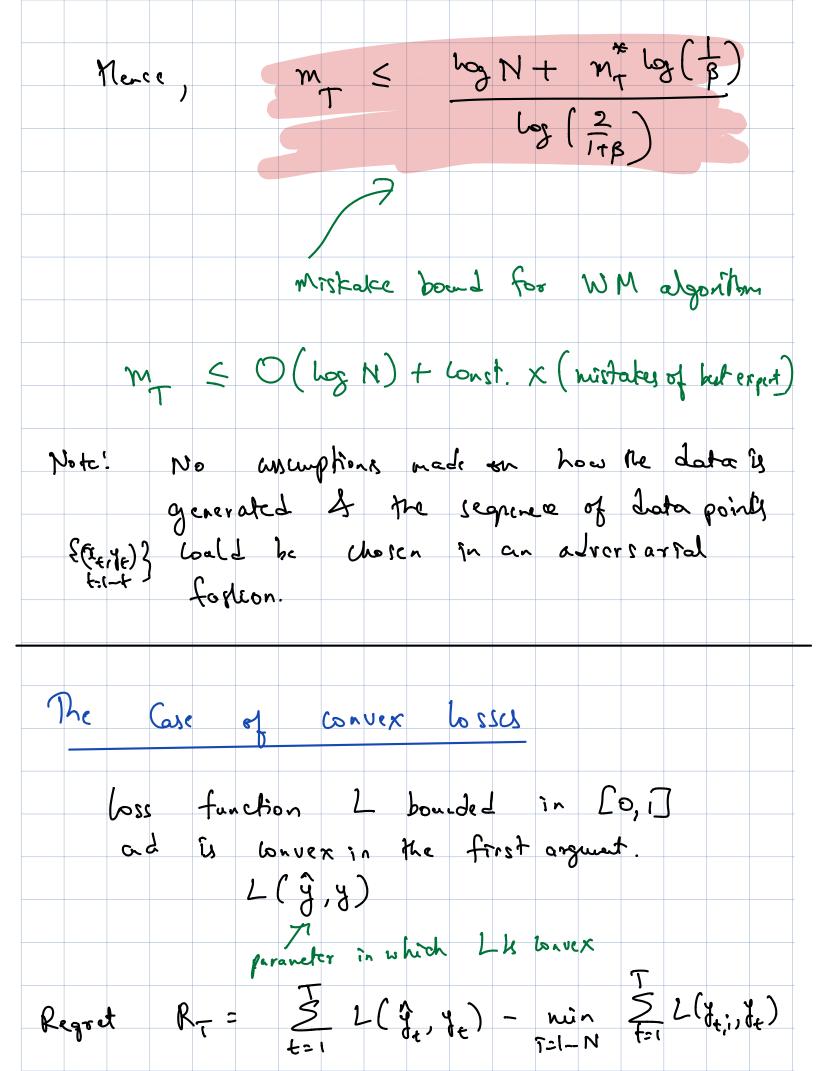
nistates done before Zerong in on the perfect enget $\mathcal{M}_{A}(c) = \max \left[\operatorname{max}_{A} \right] \operatorname{misfakes}(A, C) \right],$ $\mathcal{M}_{A} = \left[\begin{array}{c} x_{1} - x_{T} \\ x_{1} - x_{T} \end{array} \right]$ tone concept for cone T offer which them are no nuistates. For a concept closs $\frac{1}{6}$, $M_{A}(\frac{1}{6}) = \max_{c \in C} M_{A}(c)$ Halving algorithm Hove a Bet of active experts Alg A prediction! majority vote The case of noistake! remove all experts Who made a wrong prediction from the octive Set. Activo set H = H (all experts) Pseudo code! for t = 1 - T $\begin{cases} Pereive X_{t} \end{cases}$ ŷ ∈ majority vote (H_t, I_t)

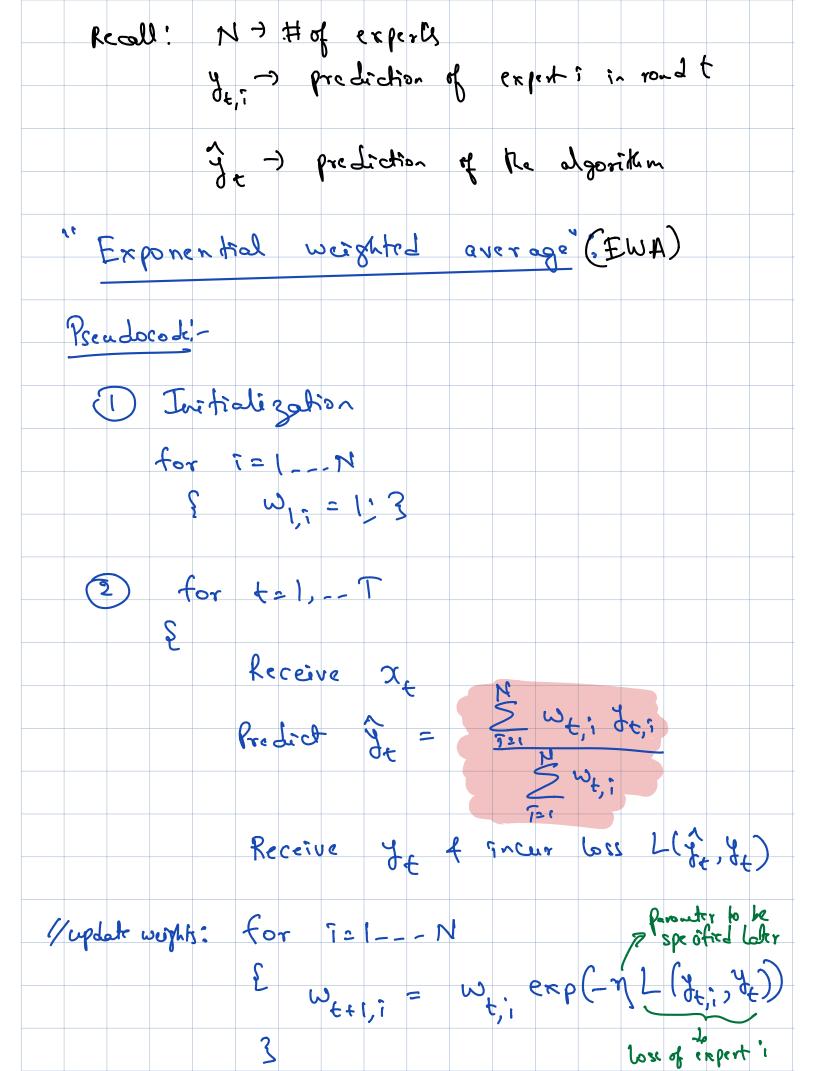
Observe y $P_{t} = \frac{1}{2} \frac{1}$ elie $H_{ffi} \subset H_{E}$ bound " Mistake $M(H) \leq \log |H|$ Proof: On a nustake, the active set is reduced by atleast half. Mence, offer by 141 mistakes, only one elevent remains in the active set 4 this can only be the perfect expert. (= target concept) Non-realizable Case & the weighted majority (WM) algorithm !



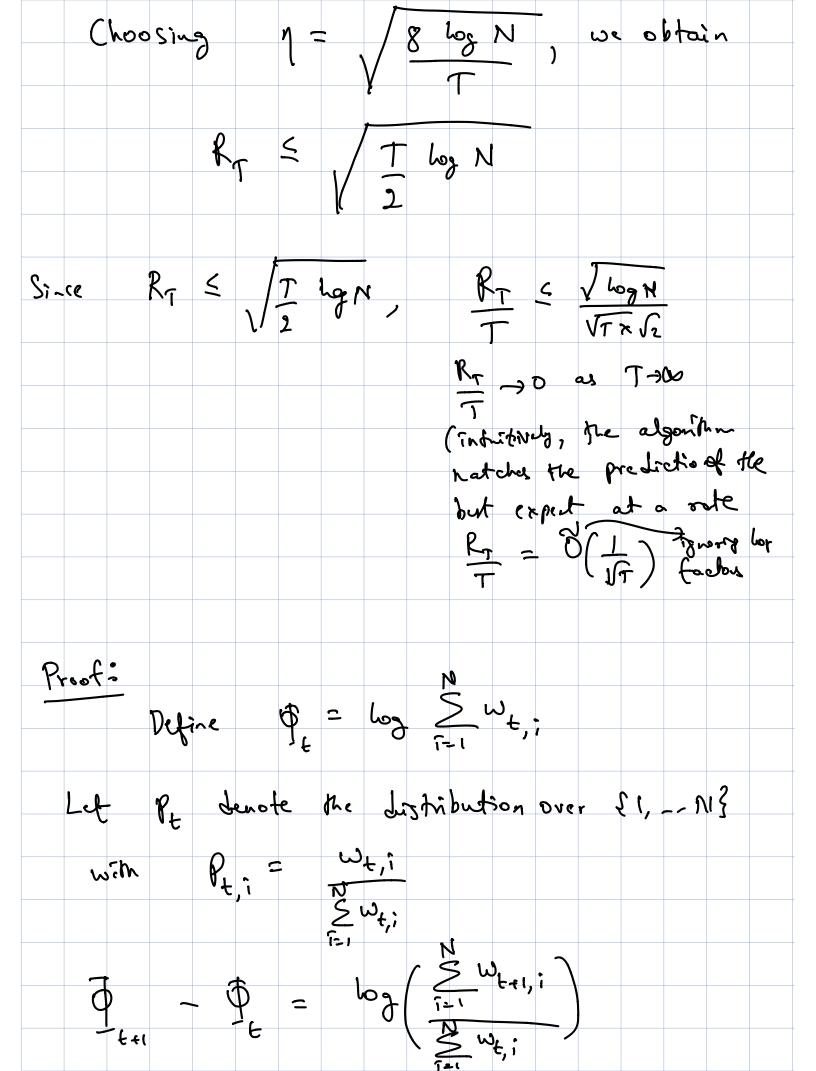


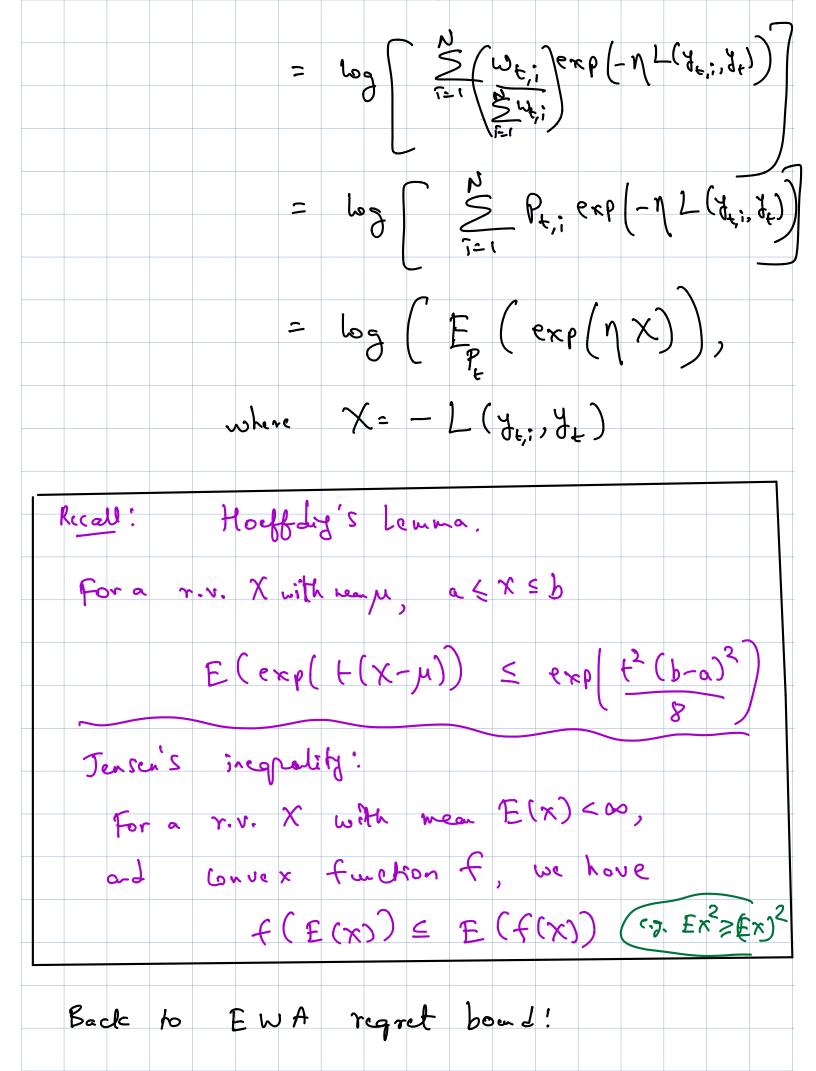


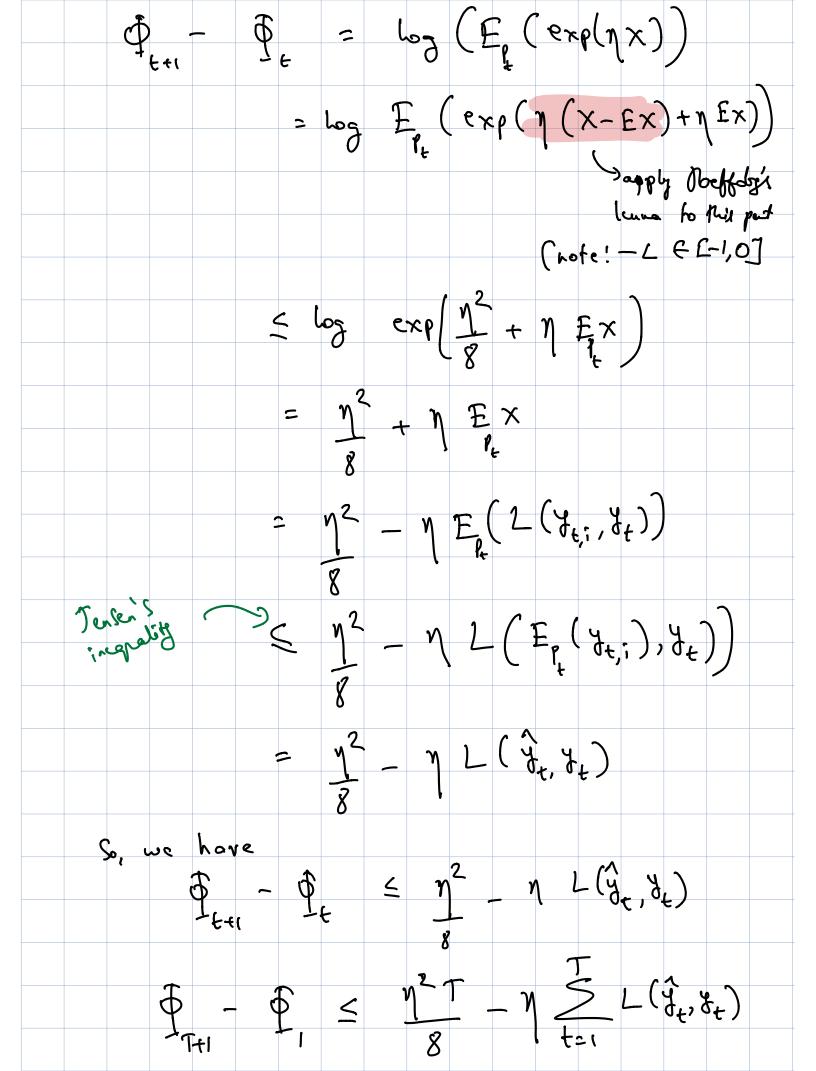


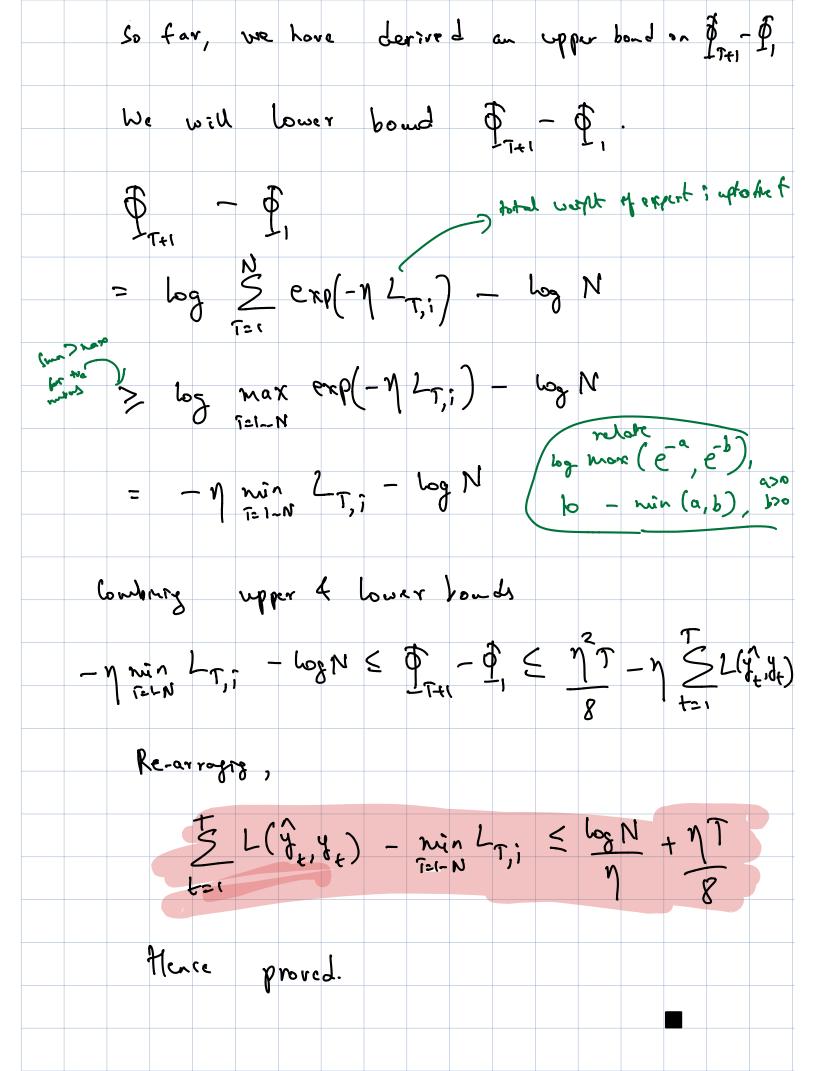


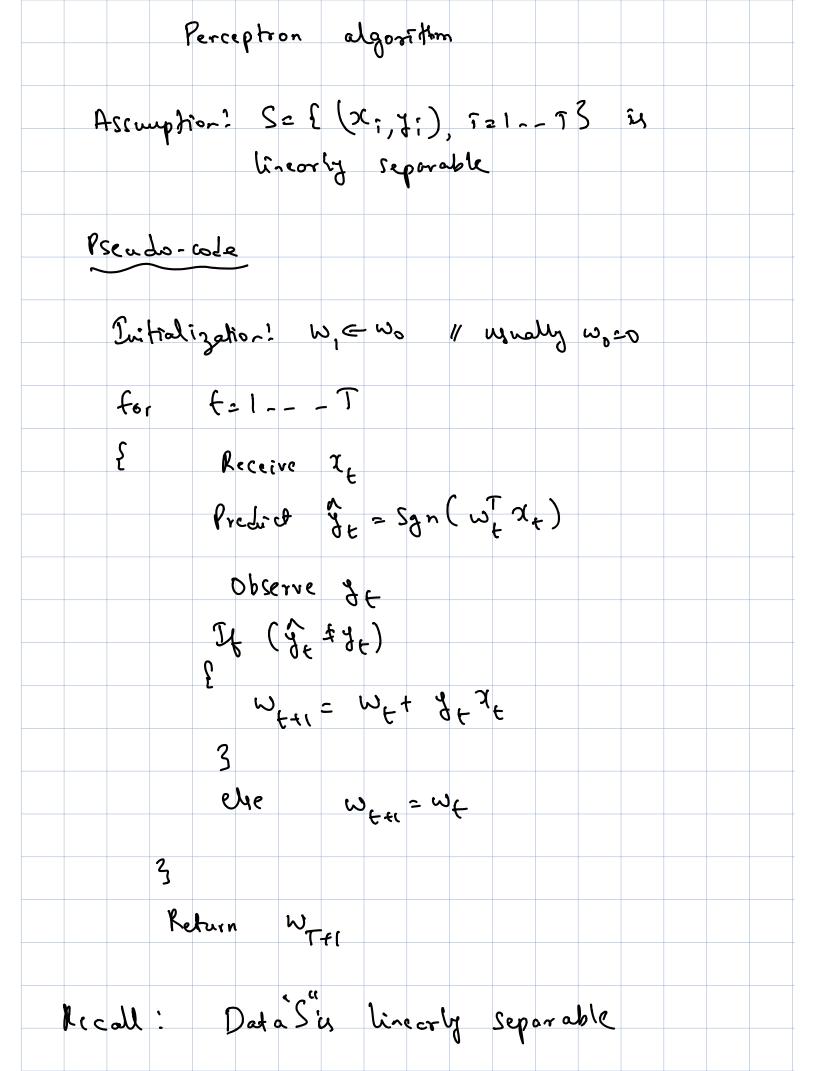
{ $W_{t+1,i} = W_{t,i} e_{x} \rho \left(-\eta L \left(\gamma_{t,i}, \gamma_{t}\right)\right)$ Nofe' $= e \kappa p \left(-\eta \frac{\xi}{\xi} L(\vartheta_{s,i}, \vartheta_{s})\right)$ Cumlative loss of expert i up to fire t = exp(-4 Lt,i) The weight of appendi on the cumlative loss A the individual losses need not be stored by the EWA algorithm Regret analysis for EWA algorithm!-Theorem ! Lis convex & bounded in Co, i] For any y>O, ad any sequence y, -- yT, the regret of EWA after Tromdy Setsifies $R_T \leq \frac{\log N}{\eta} + \frac{\eta T}{8}$

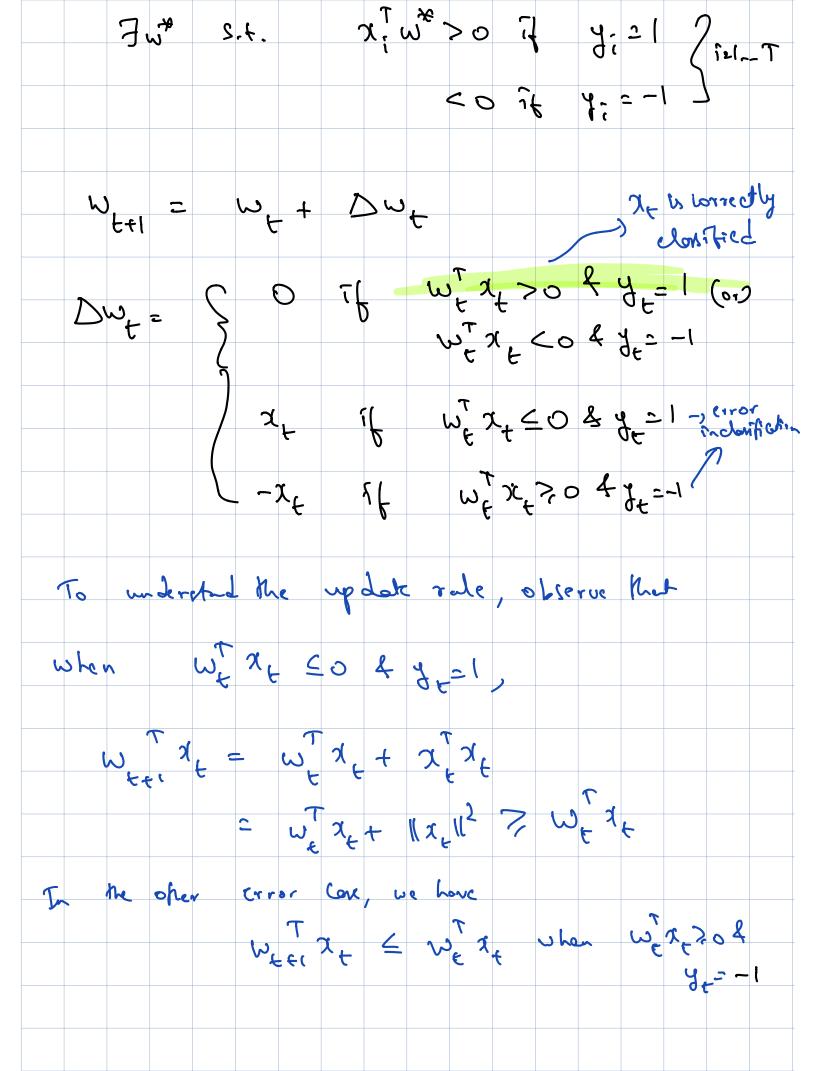


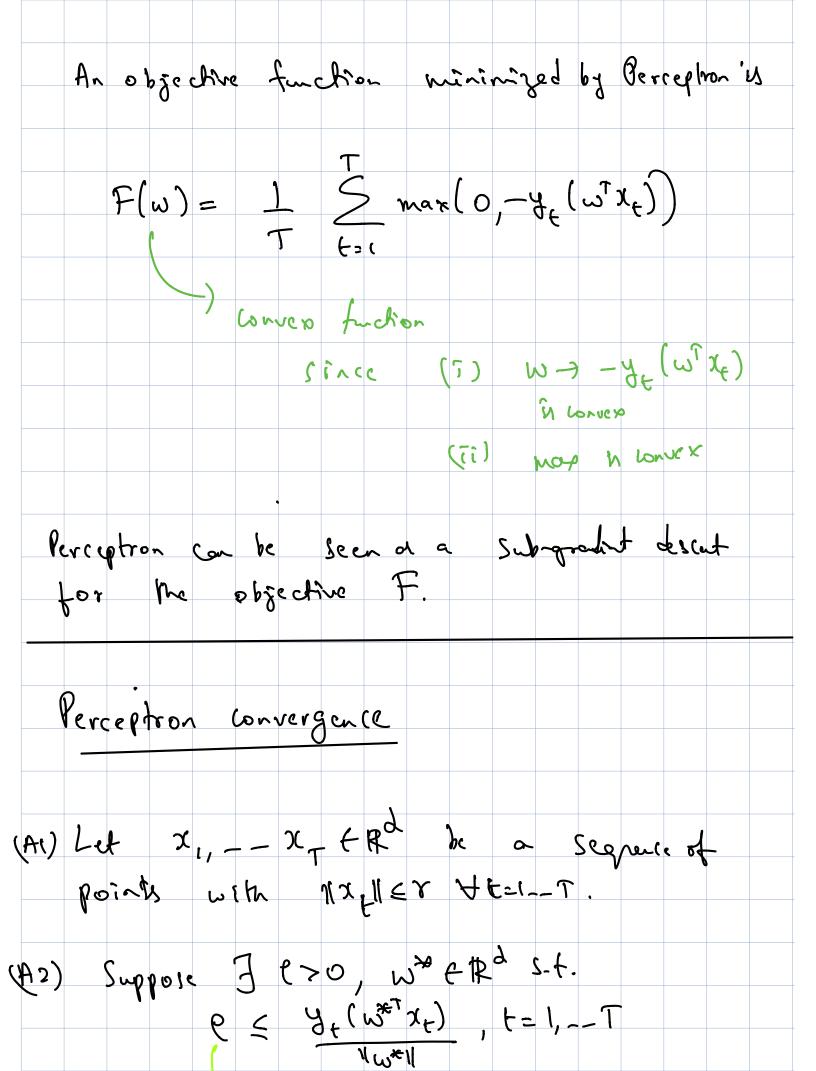


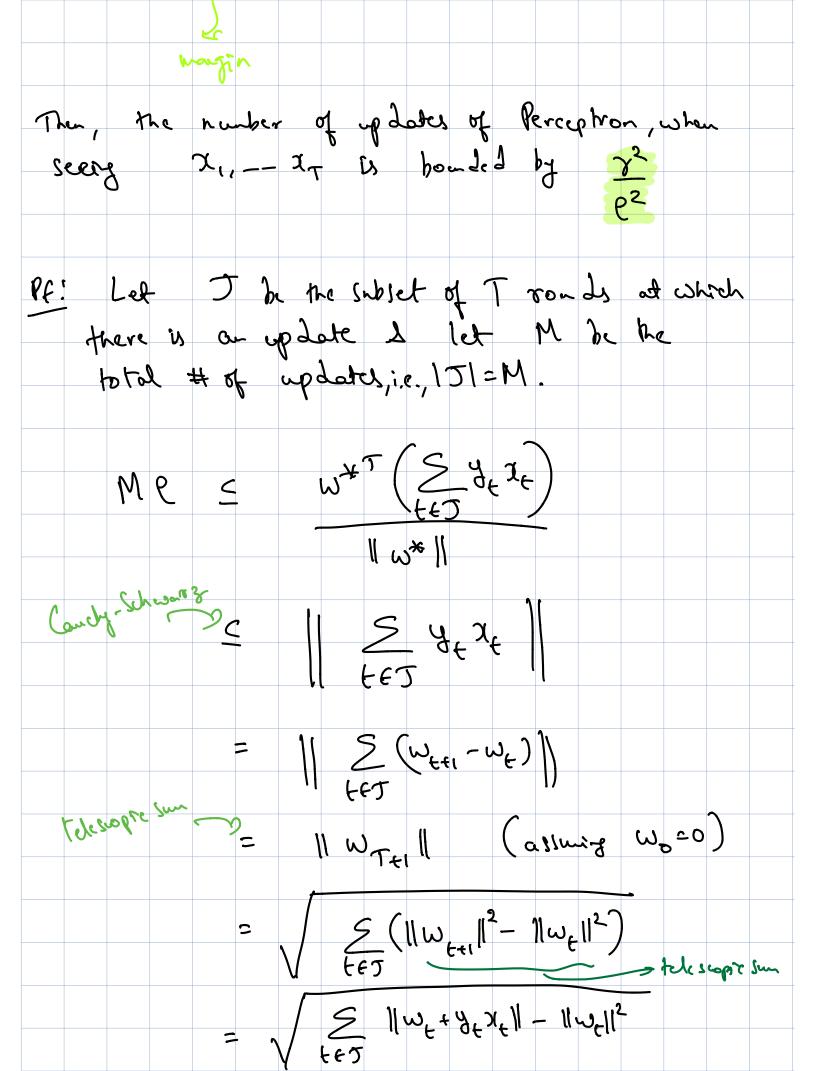












$$= \sqrt{\sum_{k \in J} 2y_{k} w_{k}^{T} x_{k} + Nx_{k} N^{k}}$$

$$= \sqrt{\sum_{k \in J} 1x_{k} N^{2}} \qquad She \quad y_{k} w_{k}^{T} x_{k} \leq 0$$

$$= \sqrt{\sum_{k \in J} 1x_{k} N^{2}} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$So, \quad \sqrt{M} \leq \overline{T} \quad or$$

$$e$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$So, \quad \sqrt{M} \leq \overline{T} \quad or$$

$$e$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$So, \quad \sqrt{M} \leq \overline{T} \quad or$$

$$e$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$So, \quad \sqrt{M} \leq \overline{T} \quad or$$

$$E$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) \leq 15 \sum_{k \in J} M$$

$$= \sqrt{M + 2} \qquad h_{j} (A_{k}) = \sqrt{M + 2} \qquad h_{j$$

Reading assignments Kernel perceptron. (figure 8.97 Forc book) Predict: $\hat{y}_{t} = \operatorname{Sgn}\left(\sum_{s=1}^{t} 2_{s} y_{s} k(x_{s}, x_{t})\right)$ Check he update reals for de